ACTA BIOLÓGICA COLOMBIANA
http://www.revistas.unal.edu.co/index.php/actabiol

Facultad de Ciencias
Departamento de Biología
Sede Bogotá

UNIVERSIDAD
NACIONAL
DE COLOMBIA

ARTÍCULO DE INVESTIGACIÓN / RESEARCH ARTICLE

IMPORTANCE OF CHARACTERISTICS IN WHITE AÇAÍ (*Euterpe oleracea* Mart.) PROGENIES BY MULTIVARIATE ANALYSIS

Importancia de las características en progenies de açaí blanco (*Euterpe oleracea* Mart.) por análisis multivariado

Gilberto Ken Iti YOKOMIZO[1*], **Ana Késsia Freitas de MATOS**[2], **Kuang HONGYU**[3], **Maria do Socorro Padilha de OLIVEIRA**[4], **Eneas Correa dos SANTOS**[5]

[1] Embrapa Amapá, Rod. Josmar Chaves Pinto, 2600, Bairro Universidade, CEP 68903-419, Amapá, Brasil, e-mail gilberto.yokomizo@embrapa.br
[2] Universidade do Estado do Amapá, Avenida Presidente Vargas, 650, Bairro Central, CEP 68900-070, Macapá, Brasil, e-mail: freitasdematosana@gmail.com
[3] Universidade Federal do Mato Grosso, Avenida Fernando Corrêa da Costa, 2367, Bairro Boa Esperança, CEP 78060-900, Cuiabá, Brasil, e-mail kuang_hongyu@hotmail.com
[4] Embrapa Amazônia Oriental, Travessa Dr, Enéas Pinheiro, s/n, Bairro Marco, CEP 66095-903, Belém, Brasil, e-mail socorro-padilha.oliveira@embrapa.br
[5] Universidade Federal Rural da Amazônia, Avenida Presidente Tancredo Neves, 2501, Bairro Terra Firme, CEP 66077-830, Belém, Brasil, e-mail: eneas_agronomo@hotmail.com
* For correspondence

Received: 23rd September 2022. **Returned for revision:** 04th January 2023. **Accepted:** 20th September 2023
Associate Editor: Magdi T. Abdelhamid

Citation/ citar este artículo como: Yokomizo, G.K., Matos, A.K.F., Hongyu, K., Oliveira, M.S.P., Santos, E.C. (2024). Importance of Characteristics in White Açaí (*Euterpe oleracea* Mart.) Progenies by Multivariate Analysis. *Acta Biol Colomb.*, *29*(2): 62-72. https://doi.org/10.15446/abc.v29n2.104524

ABSTRACT

The açaí (*Euterpe oleracea* Mart.) has an ethno-variety white characterized by the absence of anthocyanin in its epicarp, differing from violet açai. White ethno-variety is under strong anthropic pressure, which can lead to genetic erosion. Studies with white açaí that allow the selection of possible cultivars to stimulate their cultivation are scarce. Therefore, this work aimed to verify which characteristics best discriminate selected progenies of white açaí. The experiment contains 52 progenies installed, in a completely randomized design, with ten replications of one plant per plot in a lowland area belonging to Embrapa Amazônia Oriental. The characteristics evaluated were total bunch weight (PTC) in kilogram; fruit weight per bunch (PFC) in kilogram; fruit yield per bunch (RFC) in %; one hundred fruit weight (PCF) in grams; rachillas number per bunch (NRC); and rachilla length in the bunch (CRC), in centimeter. The conclusions are that there exists a correlation between the characteristics in different intensities through multivariate analysis; the Box Plot indicates that the plants respond to the annual water quantity specifically in the characteristics PTC, CRC, and PFC; for the ADC, the characteristics with the greatest contribution to the variance are PTC, CRC and PFC, being similar for the ACP mainly in the PTC and PFC characteristics; by the canonical Biplot the characteristics PTC, PFC, CRC and NRC are preferentially evaluated, discarding RFC and PCF, reducing the effort to be effective.

Keywords: ethnic group, graphic analysis, productivity, vegetal resource.

RESUMEN

El açaí (*Euterpe olearecea* Mart.) tiene una etnovariedad blanca caracterizada por la ausencia de antocianina en su epicarpio, a diferencia del açaí violeta. La etnovariedad blanca se encuentra bajo fuerte presión antrópica, lo que puede generar erosión genética. Son escasos los estudios con açaí blanco que permitan la selección de posibles cultivares para estimular su cultivo. Por lo tanto, el objetivo de este trabajo fue verificar qué características discriminan mejor las progenies seleccionadas de açaí blanco. El experimento contiene 52 progenies, en un diseño completamente al azar, diez repeticiones de una planta por parcela en un área de llanura inundable perteneciente a Embrapa Amazônia Oriental. Las características evaluadas fueron peso total del racimo (PTC) en kilogramos; peso de fruto por racimo (PFC) en kilogramos; rendimiento de fruto por racimo (RFC) en %; peso de cien frutos (PCF) en gramos; número de rachillas por racimo (NRC); y longitud de la raquila en el racimo (CRC), en centímetros. Las conclusiones, mediante el análisis multivariado, son que existe correlación entre las características en diferentes intensidades; el Box Plot indica que las plantas responden a la cantidad de agua anual específicamente en las características PTC, CRC y PFC; para la ADC las características con mayor contribución a la varianza son PTC, CRC y PFC, siendo similares para la ACP principalmente en las características PTC y PFC; por el Biplot canónico se evalúan preferentemente las características PTC, PFC, CRC y NRC, descartando RFC y PCF, reduciendo el esfuerzo para ser efectivo.

Palabras clave: grupo étnico, análisis gráfico, productividad, recurso vegetal.

INTRODUCTION

The açaí (*Euterpe oleracea* Mart.) is a palm of the Arecaceae family, widely disseminated and cultivated in the Brazilian Amazon. It occurs naturally in the Amazon region and has great economic importance for regional fruit production, especially for Pará, whose pulp production, commercialization and consumption generate large markets (Neves et al., 2015), with açaí juice being an iconic regional product of this state, representing the main product extracted from the palm tree (Silvestre et al., 2016).

With the popularization of the ecological appeal and the discoveries of nutritional qualities, the açaí pulp commercialization has gradually grown along with an expansion to other national markets and reaching international levels (Pagliarussi, 2010). This production makes Brazil the largest açaí pulp producer, consumer, and exporter (Menezes, 2005), which produced in 2020 the amount close to 1500 thousand t, in an area of 221 thousand ha, generating about R$ 4754 billion, with the North region representing 99 % of these values, highlighting Pará with more than 96 % of the total Brazilian (Instituto Brasileiro de Geografia e Estatística [IBGE], 2022).

An inherent aspect of man is that over time, traditional communities in the process of domestication of different species selected a wide intraspecific diversity. The result of this process is reflected in the large number of existing varieties, which are generally considered a cultural artifact of these communities and can therefore be called ethno-varieties (Peroni and Martins, 2000), in this way, it can be considered that the açaí tree is spread throughout the Amazon region and fits perfectly into this denomination.

The açaí tree fits as an ethno-variety, where there is a diversity of types for various morphological characters. Jardim (2000) and Pimentel and Jardim (2009) defined different types or ethno-varieties of açaí trees, based on differentiation in morphological structures, such as fruit color, the shape of inflorescences, bunches, and others, naming them as black açaí, tinga açaí, spotted açaí, mulatto açaí, sword açaí, and white açaí, additionally Jardim and Oliveira (2014) add chumbinho açaí. With white açaí being, the second most consumed, behind black açaí with almost the entire volume. It is worth noting that the fruits when ripe have a color ranging from dark purple to black, except the white type açaí, which produces light green fruits (Nogueira et al., 2005; Oliveira and Tavares, 2016).

Regarding the location, it is noteworthy that in some areas of the Marajó Island and the estuary of the Amazon River, one can find the white açaí or açaí-tinga (Oliveira et al., 2002), sword açaí on the island of Combu and in the municipality of Acará and bovine blood açaí in the southern part of Amazonas and Santarém (Simonian, 2014).

Recently, with the growth of açaí demand beyond the regional market in the North of the country, açaí planting and domestication began, to meet this growing market (Homma, 2014). Despite the increase in açaí production, stimulated by technological advances, cultivation practices, and high prices, the environmental conservation and the social return of populations have not evolved at the same pace (Teixeira, 2018). The expansion without due environmental care in this way can become an obstacle to biodiversity, by promoting the felling of other species (reduction of competition). Inserting in this context the white açaí plants loss, transforming the forest into homogeneous extensions of black açaí palms (Homma et al., 2006), that is, management practices are failing to achieve a balance between productivity and environmental conservation, corroborated by the records of Homma (2014).

Additionally, Tagore et al. (2018) affirm that the increase in market value and the amount consumed has generated the adoption of intensive and aggressive management practices for the predominance of açaí trees with black fruits. This pressure generates risk to the natural environment by stimulating the elimination of other species and, consequently, the local landscape alteration, which changes the soil and climate conditions in these regions. On the opposite side, the traditional extractive production system does not have the adoption of a cultivation process. Production occurs naturally and respects local heterogeneity. The product, in general, is intended for the maintenance of families and, therefore, is suitable as a sustainable activity with lower risks (Tagore, 2017), preserving the species present and thus avoiding the genetic erosion of white açaí that could be present.

A complicating factor in research with açaí trees in terms of selection is the evaluation of many characteristics and the effect that the environment has on the behavior of genetic material. Variations due to the response to environmental factors generate changes in its phenotypic aspects, a phenomenon known as genotypes versus environments interaction (GxA), making it difficult to select and recommend adapted and stable genotypes (Cruz et al., 2014). An important option is the use of multivariate statistics, which makes it possible to obtain the correct information and allows a joint interpretation of the characteristics under study, which are not perceptible with the use of univariate statistical analysis that works with each one in isolation (Cruz et al., 2014).

As stated earlier, almost all studies evaluated and analyzed a large number of characteristics in a univariate way, but this approach is extremely limited because it does not consider the possible correlation that may be present between the characteristics. To overcome this obstacle, multivariate analysis has been used, especially principal components analysis (PCA) and canonical discriminant analysis (CDA). The first transforms an original variables set correlated with each other into a new set of variables, called principal components, whose function is to linearly reduce the size of a data set, generating a set of p uncorrelated variables, retaining as much information as possible, holding the total variation, ordered in descending order (Hair et al., 2014; Johnson and Wichern, 2019). CDA, like ACP, reduces the dimensionality and standardizes the data. The derivation of canonical coefficients, together with a multivariate analysis of variance is the CDA base (Johnson and Wichern, 2019).

Based on the aspects of the possible loss of genetic material from white açaí, research that can stimulate and show the existence and importance of this type is essential, with their analysis not based only on the univariate dimension, but in multivariate statistical methods for the interpretation of the phenotypic variations produced by the different characteristics. Therefore, the objective of this work was to analyze the characteristics of behavior in progenies of this ethno-variety, by the multivariate and the main components analysis, to determine which attributes present the greatest contribution in the phenotypic manifestations and to verify if there is an effect of plant age on these characteristics.

MATERIALS AND METHODS

This study is part of the açaí tree genetic improvement program for fruit production in the Amazon estuary, conducting an experimental area containing a total of 52 white açaí progenies located in the Germplasm of Embrapa Amazônia Oriental, Belém, PA (1° 27’21” S and 48° 30’16” W, 10.8m). This experiment was installed in February 2003, in a completely randomized design with ten replications of one plant per plot, spaced 5 m x 5 m. The matrices were collected in municipalities of the Amazon estuary, being Curralinho (01°48’49” S and 49°47’43” W), Breves (01°40’56” S and 50°28’49” W), São Sebastião da Boa Vista (01°43’03 ”S and 49°32’27” W), Muaná (1°53’51” S and 49°23’71” W), Limoeiro do Ajurú (01°53’42” S and 49°22’51” W) and Ponta de Pedras (1°39’28” S e 48°87’05” W), located in the Marajó Archipelago, at the mouth of the Amazon River (Pará state, Brazil).

For the morphoagronomic characterization of the accessions, carried out between 2009 and 2018 were used, total bunch weight (PTC) in kg; fruit weight per bunch (PFC) in kg; fruit yield per bunch (RFC), in %, obtained by dividing PFC and PTC, multiplied by 100; one hundred fruits weight (PCF) in grams; rachillas number per bunch (NRC); and rachilla length in the bunch (CRC), expressed in cm.

In the collected data initially, the univariate analysis of variance (ANOVA) to verify the existence of differences that allow performing the other analyses. A multivariate analysis was carried out, allowing observing and understanding of the differences in behavior to the six characteristics simultaneously. The principal component analysis (PCA) and canonical discriminant analysis (CDA) were the multivariate analysis methodologies employed.

Principal component analysis

Principal component analysis (PCA) was estimated using Pearson’s correlation matrix, which corresponds to each variable having been standardized (mean equal to zero and variance equal to unity), due to measurement using different units. In this way, the dependence structure of the variables is provided by the correlation matrix (R). The equation $z = \frac{x-\bar{x}}{s}$s was employed to standard the values, where z is the standardized value of x, $\bar{x}$ is the overall mean, and s is the standard deviation of the variable.

To estimate the principal components is used linear combinations of the original variables, generating a matrix R with eigenvectors. The absolute value of an eigenvector associated with the eigenvalue of each component determines the level of importance of the features in a given principal component. Each eigenvector is calculated starting from an eigenvalue of the correlation matrix of the original data and the eigenvalues are related to the variance in each principal component (Rencher and Christensen, 2012; Fraga et al., 2016).

The first principal component (PC1) can explain most of the total additive genetic variance. The second principal component (PC2) manages to explain the second largest and so on until all the variance is explained. In a dataset with p variables, the eigenvector is estimated according to x’ = [x1, x2, ..., xp] which has a correlation matrix (R) with the eigenvalue-eigenvector pairs (λi, ei) for i = 1,2,...,p, where λ1 ≥ λ2 ≥...≥ λp ≥ 0 (Rencher and Christensen, 2012; Johnson and Wichern, 2019). The i-th principal component is estimated by the equation where e_{ip} refers to the i-th eigenvector and x_p to the p-th variable obtained from the original data.

According to Johnson and Wichern (2019), another method whose function is to assist in deciding the number of principal components is through a graphical representation, called scree plot, in which the numerical or relative value of each eigenvalue is located on the ordinate axis and the abscissa axis are the component axes. In a scree plot, the curve formed is similar to an elbow and the method consists of retaining the PCs located above the point where the eigenvalues have approximately the same value, in this case at the inflection point of the curve.

BIPLOT GRAPHIC ANALYSIS

The analysis with the aid of biplot plots (Gabriel, 1971) is based on the correlation matrix, having great interest in summarizing information (Johnson and Wichern, 2019).

The fact that any data matrix with the structure Y (n × p) of characteristic r can be factored according to Y = GH', where G is a matrix (n x r) and H is a matrix (p x r), both must have characteristic r is the base to an exploratory analysis by the biplot. This factoring is not unique. One way to factor Y is to choose the r columns of G, as an orthonormal basis of the column space of Y, and compute H as Y'G.

Canonical Discriminant Analysis

Canonical discriminant analysis (CDA) or canonical variables analysis is applied to the results of principal component analysis, aiming to differentiate treatments based on the most important characteristics. The first stage of the CDA verifies the assumptions and then the multivariate analysis of variance (MANOVA) is performed with verification of significant differences between treatments in the canonical functions using the multivariate Lambda tests of Wilkis, Pillai, Wilks, Hotelling-Hawley, Roy (Johnson and Wichern, 2019).

The use of CDA makes it possible to determine functions of the variables X=(X1, X2,..,Xp), which separate g groups as much as possible, using Z a linear combination of the original variables, and the mean value of Z changes from one group to another and thus generating separate groups. The approach aims to build linear combinations of Xi variables, in which the indices a = (a1, a2,.., ap) can maximize the F ratio for a one-way analysis of variance (Hair et al., 2014; Jonhson and Wichern, 2019). The linear combination of variables is given by: $Z = a_1X_1 + a_2X_2 + a_3X_3 + ... + a_pX_p$.

A biplot is structured for the first two canonical variables (CAN1 and CAN2), managing to discriminate the treatments jointly in terms of the most important characteristics in the canonical variables. All analyses were performed using the R 3.4.1 software (R Foundation, n.d.), considering α=5%.

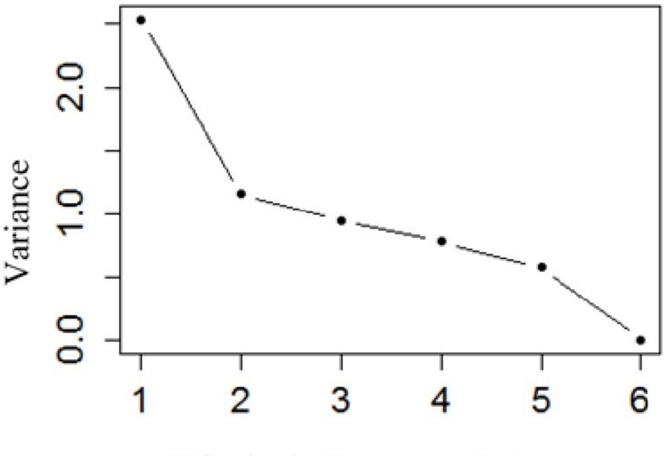

Figure 1. Scree Plot of principal component axes eigenvalues in white açaí progenies.

RESULTS

The results in Table 1, which is a preliminary univariate analysis, show statistical differences in the effect of years for all traits among the different progenies. Visualize that there are different behaviors between the progenies and thus also allowing the multivariate analysis was performed after verifying the presence of differences in univariate analysis.

In principal component analysis (PCA), the explained variance is how much each component can explain the eigenvalue, given as a percentage, and the accumulated proportion is how much each component successively accumulates from its previous ones until reaching 100 % (Araújo and Coelho, 2009). The results obtained by the principal components technique are in Table 2. The first two PCs were responsible for 61.47 % of the total variation, on the effect of years on the evaluated characteristics, where PC1 was responsible for 42.24 % and the second, PC2, for 19.24 % of the data variations.

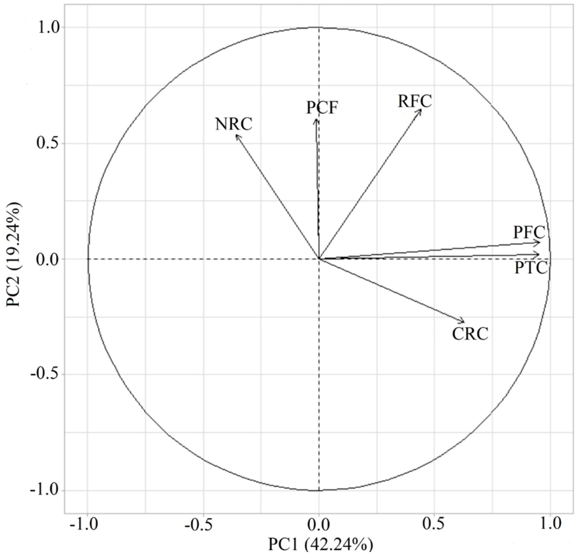

Figure 2. Eigenvectors of PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch, in the first two principal components (PC1 and PC2) in white açai progenies.

Table 1. Summary of univariate analysis of variance for PTC, PFC, RFC, CRC, NRC, and PCF in white açaí progenies.

SV	DF	MS					
		PTC	PFC	RFC	CRC	NRC	PCF
Years	9	1178.99**	803.76**	2159.26**	8390.1**	13713.6**	2284.2*
Residue	703						

PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch.

Table 2. Principal component axes (PC), eigenvalues (λi), percentage of variance explained by the components (PCV), cumulative percentage (aVPC) scores (Weights), and correlations (Correl) in the component axes (PC1, PC2 and PC3) for six characteristics evaluated in white açai progenies.

	PC1	PC2	PC3	PC1	PC2
Λi	1.5919	1.0743	0.9730		
VPC	42.24	19.24	15.78		
aVPC	42.24	61.47	77.25		
	Pesos			Correl	
	PC1	PC2		PC1	PC2
PTC	0.5986	0.0182		0.9529	0.0195
PFC	0.5992	0.0666		0.9539	0.0716
RFC	0.2765	0.6029		0.4402	0.6477
CRC	0.3946	-0.2553		0.6282	-0.2742
NRC	-0.2245	0.5006		-0.3574	0.5378
PCF	-0.0078	0.5621		-0.0124	0.6039

PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch.

The Scree Plot is a line plot of the eigenvalues of factors or principal components in evaluated charcteristics. The generate graphic is employed to determine the number of factors to retain in an exploratory factor analysis or principal components to retain in a principal component analysis (PCA) according to Matos and Rodrigues (2019).

The inflection point occurs at number two in the horizontal axis, which represents the second component axis (Fig. 1).

The graphics to components axis (Fig. 2) showed the proximity among NRC, PCF and RFC, close to the positive vertical axis and PFC, PTC close to positive horizontal axis, with CRC close to these.

For Manova, according to Song (2013), four alternative statistics were used to test the hypothesis that all samples come from populations with the same mean vector. Being considered robust (ie, they can be applied even if the assumptions do not hold for all groups or variables) if the sample sizes are approximately equal for the m samples. Thus, through the multivariate Lambda tests of Wilkis, Pillai, Hotelling-Lawley, and Roy, the F test between the different years of evaluation (Table 3) to identify the presence of significant differences.

The four different methodologies for multivariate analysis (Table 3) indicate at 1% significance that the white açaí progenies behave differently in the set of characteristics evaluated.

In (Table 4), the six canonical functions were needed to explain the 100 % of the variability of the data, but only the first two were considered since the percentage above 80 % is sufficient for reliable and satisfactory interpretation of the results, which allows the understanding of behavior in the set of parameters analyzed (Cruz et al., 2014).

In the first and second canonical component discriminant function (Table 5) the most important features (positive and/or negative) were in order, PTC, CRC, and PFC, so these variables are important in the discrimination of progenies between the years, whose absolute values were higher. The years also had different effects in phenotypic values to evaluated characteristics.

Table 3. Summary of multivariate analysis of variance, using four statistical tests, indicating the degrees of freedom of the numerator (NGL) and denominator (DGL) and the probability for the F test for the response variables (six characteristics) evaluated simultaneously.

Effect	Statistical test	Value	F value	NGL	DGL	F Prob
Years	Wilkis Lambda	0.0522	51.778	54	3563.7	**
	Pillai trace	1.7138	31.232	54	4218.0	**
	Hotelling-Lawley trace	6.5431	84.373	54	4178.0	**
	Roy's biggest root	4.8629	379.840	9	703.0	**

PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch.

Table 4. Canonical variables (CV), canonical discriminant function (Can), approximation of the F value (Fap), eigenvalues (λi), percentage of variance explained by the canonical variables (CVV) and cumulative percentage (aCVV) in the characteristics evaluated in white açai progenies.

CV	Can	Fap	Λi	CVV	aCVV
CV1	0.82944	51.778**	4.86288	74.32111	74.321
CV2	0.55701	23.809**	1.25740	19.21734	93.538
CV3	0.26089	9.758**	0.35298	5.39474	98.933
CV4	0.05569	2.685**	0.05897	0.90133	99.835
CV5	0.00772	0.760	0.00778	0.11885	99.953
CV6	0.00304	0.536	0.00305	0.04663	100.000

Table 5. Contribution of the six characteristics and the years in the first and second canonical discriminant function

			Scores		
Charac	**CAN1**	**CAN2**	**Years**	**CAN1**	**CAN2**
PTC	-5.3842	-3.4041	2009	0.0923	0.2348
PFC	-5.0902	-3.3032	2010	0.5993	0.1080
RFC	-0.3015	-0.4918	2011	1.0693	-0.3368
CRC	-5.2424	3.4224	2012	0.6828	-0.1249
NRC	4.7729	-0.9662	2013	0.9305	-0.3635
PCF	0.9163	0.2655	2014	-4.4969	4.3753
			2015	0.8248	0.0639
			2016	0.8201	-0.2770
			2017	0.7210	-0.0046
			2018	-7.7014	-2.3335

PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch.

The Box Plot of the first canonical function (Fig 3) shows that the characteristics RFC, NRC and PCF contributed positively to the behavior of the set of white açaí progenies, between the years 2008 to 2012 and 2014 to 2017. While PTC, PFC and CRC contributed to the years 2013 and 2018. This result shows different contributions of the characteristics evaluated in relation to each year of evaluation.

The Figure 4 shows the canonical biplot for the characteristics evaluated in white açaí, considering the years of study. The canonical biplot objective is to prospect the main variables responsible for generating the differentiations (González-Martín et al., 2016). In CDA the approach is direct, based on a simple visual analysis, watching the projection of the two canonical discriminant functions (Sorbolini et al., 2016).

The canonical discriminant biplot (Fig. 4) results in that the characteristics, PCF and RFC present the highest specificity of the years that were similar, except for 2013 and 2018, where CRC presents specificity to this first year mentioned and PTC and PFC to the second year. NRC differentiated itself from other characteristics.

DISCUSSION

The results in Table 1 indicate the progenies presence whose performance can be highly promising for future actions of genetic improvement, Sousa et al. (2017), also found significant differences in violet-type açaí. Multivariate analysis was performed after verifying the presence of differences in univariate analysis.

In the Table 2 the first two PCs were responsible for 61.47 % of the total variation, on the effect of years on the evaluated characteristics, where PC1 was responsible for 42.24 % and the second, PC2, for 19.24 % of the data variations, close to what informed in cassava by Gonçalves et al. (2021). Thus, the sum of these first two component axes can be considered sufficient as stated by Yang et al. (2009) in which the first two principal components must explain more than 60 % of the total variance. Therefore, in the continuation of the analyses, only these first two component axes are considered. Likewise, Venturini et al. (2013) in egg production in certain weeks found that three CPS were sufficient to explain 70.93 % of the total variance of the genetic values of the traits, even with a different species from the present study, the results obtained by these authors corroborate those found in this work.

It can be said that the portion captured by the first two axes is relevant, and it can be considered sufficient to capture the variations of genetic origin, ignoring possible noise effects, or defined also by stochastic effects, which would generate difficulties for the interpretation of the analyzes (Maia et al., 2019). Additionally, when considering that the PCA was performed on the covariance matrix of standardized variables where there is retention only of the principal components whose eigenvalue is greater than unity, it follows that only the first two principal components can be considered in the analysis (Table 2), as they were the only components with an eigenvalue greater than one.

In (Table 2), to understand the importance of each variable in the structuring of the first two components with the greatest contribution and thus the PTC and PFC characteristics stood out in the first main component, and for the second component RFC and PCF stand out.

Complementarily, in the figure of the Scree Plot, the inflection starts from the second component axis, confirming the option of retaining only these two for the analysis

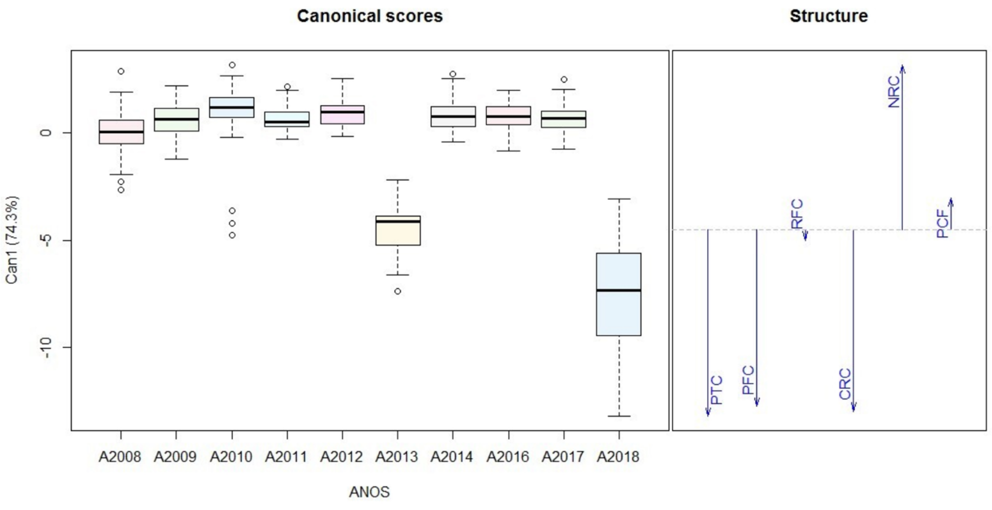

Figure 3. Boxplot plots with canonical discriminant functions for the years in the characteristics, PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch.

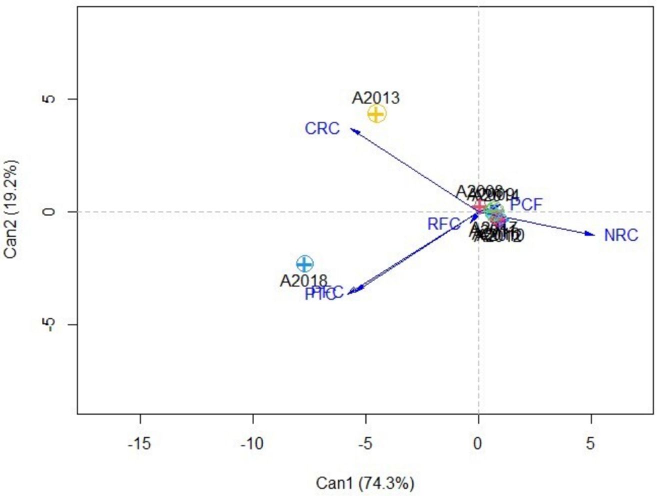

Figure 4. Canonical discriminant biplot for the correlation between treatments years and characteristics (PTC: total bunch weight; PFC: fruit weight per bunch; RFC: fruit yield per bunch; PCF: one hundred fruits weight; NRC: rachillas number per bunch and; CRC: rachilla length in the bunch) evaluated in white açai progenies.

(Figure 1), similar to those observed in Eucalyptus clones by age according to Protásio et al. (2014), that is, the most relevant information from the original sample data is contained in the first two main components.

Comparing the results observed in (Table 2) with (Fig. 2), the first component axis with the greatest contributions is those with longer vectors and closer to the component axis. In the case of the first component axis to PTC and PFC, and for the second PCF and RFC, with longer vectors and closer to the component axes, can be considered that the first axis refers more strongly to the weight characteristics inherent to the bunch and the fruits present in this bunch and, the second specifically with the fruits (RFC and PCF) more strongly associated with the interaction effects with the years of evaluation.

The PTC and PFC characteristics are more intensely correlated by the proximity between the lines, with CRC having less interaction with both. Since CRC does not correlate with NRC, PCF, and RFC, due to the angle close to or greater than 90°. The PTC and PFC traits also correlate intermediately with RFC but do not interact with NRC and PCF, while these last three traits correlate with average intensity with each other. Similar distribution and behavior of characteristics in cassava by Gonçalves et al. (2021). Similarly, in peanuts, yield traits showed greater contributions Costa et al. (2021) compared to vegetative aspects.

The results from multivariate analysis (Table 3) indicate that there is genetic divergence between the white açaí progenies considering all characteristics simultaneously, confirming the results obtained in the univariate analysis. Important result for plant selection purpose.

The percentage accumulated in the first two canonical variables (Table 4) was higher than those obtained in coffee by Teixeira et al. (2013) and like that obtained by Morais et al. (2019) in *Dalbergia ecastaphyllum*, the use of the first two variables is sufficient and safe. Representing the heritable and non-environmental behavior.

In the first canonical discriminant function (Table 5) the most important features (positive and/or negative) were in order, PTC, CRC, and PFC in contrast to the others, so this main component efficiently differentiates the clusters mainly as a function of features of bunches in terms of weight and length of rachis. The second component behaved similarly to the first when observing PTC, CRC, and PFC, so these variables are important in the discrimination of progenies between the years and should therefore be considered as the ones that researchers should concentrate greater efforts on in their field evaluations, discarding the evaluation of other characteristics, as pointed out by Teixeira et al. (2013). Regarding the years of evaluation, a significant effect of the age of the progenies can also be inferred regarding the phenotypic effect of the evaluated characteristics; similar to what was mentioned by Protásio et al. (2014), in Eucalyptus clones.

According to the Box Plot of the first canonical function, the years 2008 to 2012 and 2014 to 2017 directly influenced NRC and PCF (Fig. 3). This indicates that in these years there was a trend towards an increase in the number of rachis of the bunch and the weight of one hundred fruits. However, there was a decreasing trend for PTC, PFC, and CRC in these years. Since the Box Plot was not able to differentiate, the years mentioned above from each other, that is, possibly the intrinsic characteristics of each year were extremely similar.

Unlike in 2013 and especially 2018, environmental conditions and the age of the progenies strongly influenced the variations observed in PTC, PFC, and CRC. When consulting the climatological data, where the white açaí progenies are installed, the average, maximum, and minimum temperatures can be observed, in addition to the relative humidity from the environment in percentage, were similar throughout the evaluation period, distinguishing the years 2013 and 2018 with about 200 mm of precipitation above the average of other years (Instituto Nacional de Meteorologia [INMET], 2022). Indicating that the water availability factor presents an important factor in the productivity of açaí plants.

The canonical biplot had little similarity to the main components biplot, with the variation explained by the two main components being 61.47 % of the data variation, while the first two discriminant canonicals were able to explain 93.54 % of the total variation. Therefore, the information in Figure 4 has greater reliability. The results confirm the greater robustness of the canonical discriminant analysis compared to the PCA, in agreement with Traldi et al. (2018).

The axis of the first canonical discriminant function proved to be significantly the most important to order the effects of the years on the evaluated characteristics, as it contained 74.3 % of the observed variance. The variations existing are observed only in the years 2013 and 2018 allowed us to separate in (Fig. 4) the variables in different areas of study in the graph, being properly separated from the other years that did not contain characteristics that could generate distinction and thus were extremely similar. Note that the year 2013 directly influenced the CRC characteristic, and 2018 was more strongly associated with PTC and PFC. There is a great similarity between PTC and PFC, proximity between PCF and RFC while CRC and NRC did not show similarities in behavior with the other characteristics.

CONCLUSIONS

There is an evident correlation between the characteristics, some with greater intensity and others with less intensity, so the multivariate analysis undoubtedly has its appropriate use for the factors studied here.

By the Box Plot, the white açaí plants respond to the amount of water available annually in the production in terms of total weight of bunch and weight of fruits in bunches, specifically in the characteristics PTC, CRC, and PFC.

By the CDA, the characteristics with the greatest contribution to the variance were PTC, CRC, and PFC, being similar for the PCA mainly in the PTC and PFC characteristics.

By the canonical Biplot, the PTC, PFC, CRC, and NRC characteristics should be preferentially evaluated, and the work of evaluating RFC and PCF can be discarded, which requires a lot of manpower to be carried out.

AUTHOR'S PARTICIPATION

Conceptualization of manuscript: Yokomizo, G.K., Matos A.K.F., Hongyu, K., Oliveira, M.S.P. and Santos, E.C.

Metodology estructuration: Yokomizo, G.K., Oliveira, M.S.P. and Santos, E.C.

Software analysis: Hongyu, K.

Formal analysis: Yokomizo, G.K.

Investigation in experimenral area: Oliveira, M.S.P.

Text writing: Yokomizo, G.K., Matos A.K.F., Hongyu, K., Oliveira, M.S.P. and Santos, E.C.

Results discussion: Yokomizo, G.K., Matos A.K.F., Hongyu, K., Oliveira, M.S.P. and Santos, E.C.

Writing-review: Yokomizo, G.K.

all authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGEMENT

To Embrapa Amapá and the National Council for Scientific and Technological Development (CNPq) for granting the scholarship.

CONFLICT OF INTEREST

The authors declare no conflict of interest

REFERENCES

Araujo, W. O. and Coelho, C. J. (2009). *Análise de Componentes Principais (PCA)*. Centro Universitário de Anápolis. http://www.unievangelica.edu.br/gc/imagens/file/mestrados/artigos/RTINF_003092.pdf

Costa, J. A., Carvalho, M. A. P., Peixoto, C. P., Dantas, A. C. V. L., Almeida, A. T. and Oliveira, E. R. (2021). Divergência genética em genótipos de amendoinzeiro do grupo vegetativo valência. In: Oliveira, R. J. (Org.). *Extensão Rural: Práticas e Pesquisas para o fortalecimento da Agricultura Familiar* (254-267). Editora Científica Digital. https://doi.org/10.37885/201202455

Cruz, C. D., Carneiro, P. C. S. and Regazzi, A. J. (2014). *Modelos biométricos aplicados ao melhoramento genético*. Editora UFV.

Fraga, A. B., Silva, F. L., Hongyu, K., Santos, D. D. S., Murphy, T. W. and Lopes, F. B. (2016). Multivariate analysis to evaluate genetic groups and production traits of crossbred Holstein × Zebu cows. *Tropical Animal Health and Production*, *48*(3), 533-538. https://doi.org/10.1007/s11250-015-0985-2

Gabriel, K. R. (1971). The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. *Biometrika*, *58*(3), 453-467. https://doi.org/10.1093/biomet/58.3.453

Gonçalves, Z. S., Lima, L. K. S., Borges, C. V. and Rocha, A. J. (2021). Avaliação agronômica e qualidade de farinha em cultivares de mandioca sob condições de campo. *Journal of Biotechnology and Biodiversity*, *9*(2), 192-200. https://doi.org/10.20873/jbb.uft.cemaf.v9n2.goncalves

González-Martín, M. I., Vicente-Tavera, S., Revilla, I., Vivar-Quintana, A. M., González-Perez, C., Hernández Hierro, J. M. and Lobos-Ortega, I. (2016). The role of the canonical biplot method in the study of volatile compounds in cheeses of variable composition. *Grasas y Aceites*, *67*(1), 1-8. https://doi.org/10.3989/gya.0250151

Hair, J. F., Black, W. C., Barry, B. J., Babin, B. J. and Anderson, R. E. (2014). *Multivariate Data Analysis*. Pearson Education Limited.

Homma, A. K. (2014). Extrativismo vegetal ou plantio: qual a opção para a Amazônia? In: Homma, A. K. O. (Org.). *Extrativismo vegetal na Amazônia: história, ecologia, economia e domesticação* (17-43). Embrapa.

Homma, A. K. O., Nicoli, C. M. L., Menezes, A. J. E. A., Matos, G. B., Carvalho, J. E. U. and Nogueira, O. L. (2006). *Custo operacional de açaizeiro irrigado no nordeste paraense*. (Documentos, 255), Embrapa Amazônia Oriental.

Instituto Brasileiro de Geografia e Estatística [IBGE] (2020). *Produção de açaí (cultivo), 2020*. https://www.ibge.gov.br/explica/producao-agropecuaria/acai-cultivo/br

Instituto Nacional de Meteorologia [INMET] (2022). *Banco de Dados Meteorológicos do INMET*. https://bdmep.inmet.gov.br/#

Jardim, M. A. G. (2000). *Morfologia e ecologia do açaizeiro Euterpe oleracea Mart. e das etnovariedades espada e branco em ambientes de várzea do estuário*. [Tese de doutorado em Ciências Biológicas, Universidade Federal do Pará/UFPA].

Jardim, M. A. G. and Oliveira, F. G. (2014). Morfologia Floral de Duas Etnovariedades de Euterpe oleracea Mart. do Estado do Pará. *Biota Amazônia*, *4*(4), 6-9. https://doi.org/10.18561/2179-5746/biotaamazonia.v4n4p6-9

Johnson, R. A. and Wichern, D. W. (2019). *Applied multivariate statistical analysis.* Pearson.

Maia, M. C. C., Almeida, A. S., Araújo, L. B., Dias, C. T. S., Oliveira, L. C., Yokomizo, G. K. I., Rosado, R. D. S., Cruz, C. D., Vasconcelos, L. F. L., Lima, P. S. C. and Medina-Macedo, L. (2019). Principal component and biplot analysis in the agro-industrial characteristics of *Anacardium* spp. *European Scientific Journal*, *15*(30), 21-31. https://doi.org/10.19044/esj.2019.v15n30p21

Matos, D. A. S. and Rodrigues, E. C. (2019). *Análise fatorial*. ENAP.

Menezes, E. M. S. (2005). *Efeito da alta pressão hidrostática em polpa de açaí pré-congelada (Euterpe oleracea, Mart.)*. [Dissertação de Mestrado–Ciência e Tecnologia de Alimentos, Universidade Federal Rural do Rio de Janeiro/UFRRJ]. http://livros01.livrosgratis.com.br/cp001274.pdf

Morais, D. V., Nunes, L. A., Mata, V. P., Costa, M. A. P. C., Sodré, G. S. and Carvalho, C. A. L. (2019). Leaf geometric morphometrics among populations of Dalbergia ecastaphyllum (L.) Taub. *Bioscience Journal*, *35*(6), 1789-1798. https://doi.org/10.14393/BJ-v35n6a2019-3981

Neves, L. T. B. C., Campos, D. C. S., Mendes, J. K. S., Urnhani, C. O. and Araújo, K. G. M. (2015). Quality of fruits manually processed of açaí (*Euterpe oleracea* Mart.) and bacaba (*Oenocarpus bacaba* Mart.). *Revista Brasileira de Fruticultura*, *37*(3), 729-738. https://doi.org/10.1590/0100-2945-148/14

Nogueira, O. L., Figueiredo, F. J. C. and Müller, A. A. (2005). *Açaí. Belém: Embrapa Amazônia Oriental*. (Sistemas de produção, 4). Embrapa Amazônia Oriental.

Oliveira, L. P. de and Tavares, G. S (2016). *Programa de Desenvolvimento da Cadeia Produtiva do Açaí no Estado do Pará–PROAÇAÍ – PA* (41p.), SEDAP.

Oliveira, M do S. P de, Carvalho, J. E. U., Nascimento, W. M. O. and Müller, C. H. (2002). *Cultivo do açaizeiro para produção de frutos. Belém: Embrapa Amazônia Oriental*. (Circular técnica, 26). Embrapa Amazônia Oriental.

Pagliarussi, M. S. (2010). *A cadeia produtiva agroindustrial do açaí: estudo da cadeia e proposta de um modelo matemático*. [Trabalho de Conclusão de Curso, Universidade de São Paulo]. http://www.tcc.sc.usp.br/tce/disponiveis/18/180830/tce-19012011-160154/publico/Pagliarussi_Marina_Sanches.pdf

Peroni, N. and Martins, P. S. (2000). Influência da dinâmica itinerante na geração de diversidade de etnovariedades cultivadas vegetativamente. *Interciência*, *25*(1), 22-29. https://www.redalyc.org/articulo.oa?id=33904304

Pimentel, M. S. and Jardim, M. A. G. (2009). Morfologia das infrutescências e inflorescências do açaí-preto (*Euterpe oleracea* Mart.) e das etnovariedades branco, espada e tinga (*Euterpe* spp.). In: Jardim, M. A. G. (Org.). *Diversidade biológica das áreas de proteção ambiental, Ilhas do Combu e Algodoal-Maiandeua* (Coleção Adolpho Ducke), (79-87). Museu Paraense Emilio Goeldi.

Protásio, T. P., Neves, T. A., Reis, A. A. and Trugilho, P. F. (2014). Efeito da idade e clone na qualidade da madeira de *Eucalyptus* spp. visando à produção de bioenergia. *Ciência Florestal*, *24*(2), 465-477. https://doi.org/10.5902/1980509814587

R Foundation (n.d.) *The R Project for Statistical Computing*. http://www.R-project.org/

Rencher, A. C. and Christensen, W. F. (2012). *Methods of Multivariate Analysis*. John Wiley & Sons. https://doi.org/10.1002/9781118391686

Silvestre, W. V. D., Pinheiro, H. A., Souza, R. O. R. M. and Palheta, L. F. (2016). Morphological and physiological responses of açaí seedlings subjected to different watering regimes. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *20*(4), 364-371. https://doi.org/10.1590/1807-1929/agriambi.v20n4p364-371

Simonian. L. T. L. (2014). *Açaí, ah! Em Belém do Pará tem! Natureza, cultura e sustentabilidade*. Editora do NAEA/UFPA.

Song, F. N. (2013). *Técnicas de análise multivariada com aplicações a dados de natureza Biológica*. [Trabalho de Conclusão de Curso, Universidade Estadual Paulista–Instituto de Biociências]. http://hdl.handle.net/11449/121367

Sorbolini, S., Gaspa, G., Steri, R., Dimauro, C., Cellesi, M., Stella, A., Marras, G., Marsan, P. A., Valentini, A. and Macciota, N. P. P. (2016). Use of canonical discriminant analysis to study signatures of selection in cattle. *Genetics Selection Evolution*, *48*(1), 1-13. https://doi.org/10.1186/s12711-016-0236-7

Sousa, A. M., Oliveira, M. S. P. and Farias Neto, J. T. (2017). Genetic divergence among white-type acai palm accessions based on morpho-agronomic characters. *Pesquisa Agropecuária Brasileira*, *52*(9), 751-760. https://doi.org/10.1590/S0100-204X2017000900007

Tagore, M. P. B. (2017). *O aumento da demanda do açaí e as alterações sociais, ambientais e econômicas: o caso das várzeas de Abaetetuba, Pará*. [Dissertação de Mestrado–Gestão de Recursos Naturais e Desenvolvimento Local na Amazônia, Universidade Federal do Pará/UFPA]. http://repositorio.ufpa.br/jspui/handle/2011/9548

Tagore, M. P. B., Canto, O. and Sobrinho, M. V. (2018). Políticas públicas e riscos ambientais em áreas de várzea na Amazônia: o caso do PRONAF para produção do açai. *Desenvolvimento e Meio ambiente*, *45*, 94-214. https://doi.org/10.5380/dma.v45i0.51585

Teixeira, A. L., Gonçalves, F. M. A., Rezende, J. C. de, Rocha, R. B. y Pereira, A. A. (2013). Análise de componentes principais em caracteres morfológicos de café arábica em estádio juvenil. *Coffee Science*, *8*(2), 205-210. http://www.sbicafe.ufv.br:80/handle/123456789/7971

Teixeira, I. L. S. (2018). *Potencial produtivo e econômico do açaí (Euterpe Oleracea Mart.) no estado do Pará*. [Dissertação de Mestrado, Departamento de Cartografia, Universidade Federal de Minas Gerais/UFMG]. https://repositorio.ufmg.br/handle/1843/EQVA-BBWE4U

Traldi, A. B., Silva, F. L., Hongyu, K., Sartorio, S. D. and Menten, J. F. M. (2018). Características dos componentes de ovos da linhagem Ross de idades diferentes via análise multivariada. *Revista de Ciências Agrárias*, *41*(2), 557-566. https://doi.org/10.19084/RCA17145

Venturini, G. C., Savegnago, R. P., Nunes, B. N., Ledur, M. C., Schmidt, G. S., El Faro, L. and Munari, D. P. (2013). Genetic parameters and principal component analysis for egg production from White Leghorn hens. *Poultry Science*, *92*(9), 2283–2289. https://doi.org/10.3382/ps.2013-0312

Yang, R. C., Crossa, J., Cornelius, P. L. and Burgueño, J. (2009). Biplot analysis of genotype x environment interaction: Proceed with caution. *Crop Science*, *49*(5), 1564-1576. https://doi.org/10.2135/cropsci2008.11.0665